



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 950345

TO: Jeanine Goldberg
Location: mail 12E12; room 12D11
Art Unit: 1634
Tuesday, May 27, 2003

Case Serial Number: 09/935464

From: Barb O'Bryen
Location: Biotech-Chem Library
CM1-6A05
Phone: 308-4291

BOS

barbara.obryen@uspto.gov

Seal

O'Bryen, Barbara

From: Goldberg, Jeanine
Sent: Tuesday, May 20, 2003 6:14 AM
To: O'Bryen, Barbara
Subject: RE: 09/935,464- schizophrenia

Thank you. I called applicant.

1. Please search SEQ ID NO: 12-13 and 39 only.

THanks

-----Original Message-----

From: O'Bryen, Barbara
Sent: Monday, May 19, 2003 4:42 PM
To: Goldberg, Jeanine
Subject: RE: 09/935,464- schizophrenia

Hi Jeanine,
this case has only 90 seqs. How would you like to modify this request?
Barb

-----Original Message-----

From: Goldberg, Jeanine
Sent: Monday, May 19, 2003 3:36 PM
To: O'Bryen, Barbara
Subject: 09/935,464- schizophrenia

1. please search SEQ ID NO: 39, 12-13, 89-92, 100, 101.

THANK YOU
Jeanine

Jeanine Enewold Goldberg
1634
CM1--12D11
Mailbox-- 12E12
306-5817



Gencore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2003, 09:39:57 ; Search time 705.6 Seconds

(without alignments)
866.155 Million cell updates/sec

Title: US-09-935-464-13

Perfect score: 21

Sequence: 1 accttccctgcagacctgttc 21

Scoring table:

IDENTITY_NDC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

GenBank: 1: gb_ba: 2: gb_hg: 3: gb_in: 4: gb_cm: 5: gb_ov: 6: gb_pat: 7: gb_ph: 8: gb_pl: 9: gb_pr: 10: gb_ro: 11: gb_sts: 12: gb_sy: 13: gb_un: 14: gb_vi: 15: gb_vl: 16: em_fun: 17: em_hum: 18: em_in: 19: em_mu: 20: em_cm: 21: em_or: 22: em_ov: 23: em_pat: 24: em_ph: 25: em_pl: 26: em_ro: 27: em_sts: 28: em_un: 29: em_vi: 30: em_hg_hum: 31: em_hg_inv: 32: em_hg_other: 33: em_hg_mus: 34: em_hg_pin: 35: em_hg_red: 36: em_hg_mam: 37: em_hg_vrt: 38: em_sy: 39: em_hgo_hum: 40: em_hgo_mus: 41: em_hgo_other:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	21	100.0	157875	9 HS272L16	AL023754 Human DNA
C 2	19	90.5	192169	2 AC020849	AC020849 Mus muscu
C 3	19	90.5	206884	2 AC113858	AC113858 Mus muscu
C 4	18.4	87.6	163505	2 AC126986	AC126986 Rattus no
C 5	18.4	87.6	192938	2 AC103908	AC103908 Canis fam
C 6	18.4	87.6	198489	2 AC100820	AC100820 Canis fam
C 7	18.4	87.6	201000	2 AC105150	AC105150 Homo sapi
C 8	18.4	87.6	201973	2 AC017049	AC017049 Homo sapi
C 9	18.4	87.6	218992	2 AC103803	AC103803 Canis fam
C 10	17.8	84.8	425	5 APERDNA	X57379 A.platyrrhin
C 11	17.8	84.8	1871	9 BC006383	BC006383 Homo sapi
C 12	17.8	84.8	1923	9 AF110957	AF110957 Homo sapi
C 13	17.8	84.8	1960	9 AB006969	AB006969 Homo sapi
C 14	17.8	84.8	2050	9 AB002135	AB002135 Homo sapi
C 15	17.8	84.8	2081	9 BC003171	BC003171 Homo sapi
C 16	17.8	84.8	2108	9 BC004129	BC004129 Homo sapi
C 17	17.8	84.8	2582	9 HSM802492	AL157437 Homo sapi
C 18	17.8	84.8	2617	9 HS035832	U35832 Human anthr
C 19	17.8	84.8	2627	9 HSM801873	AL136905 Homo sapi
C 20	17.8	84.8	2648	9 AF090384	AF090384 Homo sapi
C 21	17.8	84.8	2650	9 BC003153	BC003153 Homo sapi
C 22	17.8	84.8	2663	9 AK023120	AK023120 Homo sapi
C 23	17.8	84.8	2683	9 AF079566	AF079566 Homo sapi
C 24	17.8	84.8	4097	9 AB002137	AB002137 Homo sapi
C 25	17.8	84.8	5026	5 AB089503	AB089503 Nipponia
C 26	17.8	84.8	11464	3 DR01AMB2A	M58417 Drosophila
C 27	17.8	84.8	48836	2 AC020151	AC020151 Drosophila
C 28	17.8	84.8	61213	2 AC105017	AC105017 Homo sapi
C 29	17.8	84.8	62565	2 AC102806	AC102806 Mus muscu
C 30	17.8	84.8	82975	2 AC130171	AC130171 Rattus no
C 31	17.8	84.8	87496	2 AC125497	AC125497 Gallus ga
C 32	17.8	84.8	104878	2 AC103132	AC103132 Rattus no
C 33	17.8	84.8	144749	2 AC025912	AC025912 Mus muscu
C 34	17.8	84.8	174817	2 AC127186	AC127186 Rattus no
C 35	17.8	84.8	178886	2 AC116832	AC116832 Mus muscu
C 36	17.8	84.8	181063	3 AC010107	AC010107 Drosophila
C 37	17.8	84.8	184657	3 AC010043	AC010043 Drosophila
C 38	17.8	84.8	189495	10 AL607083	AL607083 Mouse DNA
C 39	17.8	84.8	194531	2 AC124924	AC124924 Rattus no
C 40	17.8	84.8	285978	2 AE003551	AE003551 Drosophila
C 41	17.4	82.9	43883	2 AC004396	AC004396 pseudomon
C 42	17.4	82.9	61245	2 AC044824	AC044824 Homo sapi
C 43	17.4	82.9	92822	8 AC005917	AC005917 Arabidops
C 44	17.4	82.9	96102	2 AC11304	AC11304 Rattus no
C 45	17.4	82.9	97495	8 AC003058	AC003058 Arabidops

ALIGNMENTS

RESULT 1
HS272L16/C
LOCUS
DEFINITION
Human DNA sequence from clone 272L16 on chromosome 1q32.1-32.3.
Contains the 3' end of the LAMB3 gene for Laminin, Beta 3 (Nlcn,
Kalinin, BM600) and a novel Rat Ca2+/Calmodulin dependent protein
kinase like gene. Contains ESTs, STSS, GSSs, genomic marker D1S491
and a ca repeat polymorphism, complete sequence.
ACCESSION
AL023754.1 GI:4007152
VERSION
KBTW0RDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 157875)
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

AUTHORS Graham/D.
TITLE Direct Submission
JOURNAL Submitted (27-NOV-1998) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone
requests: clonerequests@sanger.ac.uk

COMMENT On Dec 12, 1998 this sequence version replaced gi:3873472.

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence is the entire insert of clone 272116. This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.

272116 is from the library RPc11 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/VECTOR:pcrPAC2> This sequence was generated from a human chromosome 1 bacterial clone contig, constructed in collaboration by the Sanger Centre chromosome 1 mapping group and Brian Schutte, Bryan Bjork, Kevin Coppae and Jeffrey Murray. Department of Pediatrics, University of Iowa, USA. Further information can be found at <http://www.sanger.ac.uk/MGP/chrl.>

FEATURES

source

Location/Qualifiers

1..157875

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="1"

/map="q32.1-q32.3"

/clone="RP1-272116"

/clone_11b="RPc1-1"

966..1078

repeat_region

/note="LTR11 repeat: matches 281. .410 of consensus"

1082..1506

repeat_region

/note="MSMD repeat: matches 1. .426 of consensus"

1585..1875

repeat_region

/note="LTR1A1 repeat: matches 1. .319 of consensus"

1898..2024

repeat_region

/note="LTR11 repeat: matches 58. .203 of consensus"

2347..2925

repeat_region

/note="MER34 repeat: matches 6. .543 of consensus"

3090..3417

repeat_region

/note="LTR1A2 repeat: matches 23. .374 of consensus"

3677..4074

repeat_region

/note="MER34B repeat: matches 486. .902 of consensus"

4089..5338

repeat_region

/note="LTR1A2-internal repeat: matches 358. .1643 of consensus"

5339..5532

repeat_region

/note="Alu10 repeat: matches 101. .302 of consensus"

5536..5966

repeat_region

/note="LTR1A1-internal repeat: matches 5. .450 of consensus"

5972..6019

repeat_region

/note="LTR1A1 repeat: matches 318. .365 of consensus"

6020..6459

repeat_region

/note="LTR7 repeat: matches 1. .450 of consensus"

6460..6781

repeat_region

/note="LTR1A1 repeat: matches 1. .318 of consensus"

6786..6870

repeat_region

/note="MIR repeat: matches 56. .142 of consensus"

8020..8067

repeat_region

/note="MIR repeat: matches 95. .146 of consensus"

8150..8293

repeat_region

/note="MIR repeat: matches 68. .212 of consensus"

8380..9083

repeat_region

repeat_region

/note="MER50 repeat: matches 1. .711 of consensus"

9084..9303

repeat_region

/note="MIR repeat: matches 13. .262 of consensus"

9493..9795

repeat_region

/note="ALUSX repeat: matches 1. .304 of consensus"

10108..10282

repeat_region

/note="L2 repeat: matches 2506. .2695 of consensus"

10929..11005

repeat_region

/note="L2 repeat: matches 2671. .2748 of consensus"

12092..12129

repeat_region

/note="L2 repeat: matches 2671. .2748 of consensus"

12095..12130

repeat_region

/note="9 copies 4 mer acac 86% conserved"

12722..12855

repeat_region

/note="MER91A repeat: matches 20. .186 of consensus"

13108..13221

repeat_region

/note="MIR repeat: matches 80. .191 of consensus"

complement(13563..13979)

/note="match: GSSS A0075457 A0076848"

14235..14279

repeat_region

/note="MIR repeat: matches 60. .105 of consensus"

14327..14442

repeat_region

/note="MIR repeat: matches 112. .249 of consensus"

14495..15046

repeat_region

/note="L2 repeat: matches 2127. .2750 of consensus"

16671..16859

repeat_region

/note="MER5A repeat: matches 2. .189 of consensus"

17103..17476

repeat_region

/note="L2 repeat: matches 17. .422 of consensus"

17679..17938

repeat_region

/note="MIR repeat: matches 7. .262 of consensus"

17891..17953

repeat_region

/note="L2 repeat: matches 2694. .2749 of consensus"

18550..18758

repeat_region

/note="L2 repeat: matches 2641. .2750 of consensus"

19246..19406

repeat_region

/note="MER5A repeat: matches 4. .189 of consensus"

19460..19871

repeat_region

/note="L2 repeat: matches 2309. .2748 of consensus"

19916..19987

repeat_region

/note="MIR repeat: matches 63. .140 of consensus"

20200..20228

repeat_region

/note="MER34 repeat: matches 510. .538 of consensus"

20233..20365

repeat_region

/note="MER34 repeat: matches 413. .543 of consensus"

20338..20388

repeat_region

/note="LTR29 repeat: matches 454. .503 of consensus"

20459..20809

repeat_region

/note="MER47A repeat: matches 2. .366 of consensus"

20976..21143

repeat_region

/note="MER34 repeat: matches 6. .172 of consensus"

21834..22278

repeat_region

/note="LTR1C repeat: matches 1. .466 of consensus"

22477..22528

repeat_region

/note="26 copies 2 mer ag 79% conserved"

22826..22937

repeat_region

/note="MER81 repeat: matches 1. .112 of consensus"

24897..25034

repeat_region

/note="MER5A repeat: matches 30. .111 of consensus"

25084..25193

repeat_region

/note="MIR repeat: matches 26. .145 of consensus"

25207..25435

repeat_region

/note="MIR repeat: matches 13. .259 of consensus"

25919..26363

repeat_region

/note="LTR1P repeat: matches 68. .541 of consensus"

26364..26813

repeat_region

/note="MER4A2 repeat: matches 1. .503 of consensus"

26814..26885

repeat_region

/note="LTR1F repeat: matches 1. .68 of consensus"

26931..27357

repeat_region

/note="L2 repeat: matches 2292. .2723 of consensus"

27779..27820

repeat_region

/note="21 copies 2 mer ca 100% conserved"

```
repeat_region 27779..27818
/note="10 copies 4 mer caca 100% conserved"
repeat_region 29036..29195
/note="HAL1 repeat: matches 4..163 of consensus"
repeat_region 29263..29500
/note="HAL1 repeat: matches 141..355 of consensus"
repeat_region 29501..29820
/note="MLT1A1 repeat: matches 1..365 of consensus"
repeat_region 29821..30076
/note="HAL1 repeat: matches 355..618 of consensus"
repeat_region 30095..30228
/note="FLAM-A repeat: matches 7..142 of consensus"
repeat_region 30246..30440
/note="HAL1 repeat: matches 683..874 of consensus"
repeat_region 30539..30768
/note="Charliela repeat: matches 1189..1455 of consensus"
repeat_region 30789..31099
/note="Alus4 repeat: matches 1..306 of consensus"
repeat_region 31100..31167
/note="Charliela repeat: matches 1125..1189 of consensus"
repeat_region 31170..31914
/note="LIM4 repeat: matches 5448..6190 of consensus"
repeat_region 31923..32141
/note="LIM2 repeat: matches 4317..4530 of consensus"
repeat_region 32142..32551
/note="MSTA repeat: matches 5..426 of consensus"
repeat_region 32552..33436
/note="LIM2 repeat: matches 3451..4317 of consensus"
repeat_region 33437..33756
/note="AluY repeat: matches 1..309 of consensus"
repeat_region 33757..36041
/note="LIM2 repeat: matches 761..3451 of consensus"
repeat_region 37817..38850
/note="Charliela repeat: matches 1..1142 of consensus"
repeat_region 39293..39437
/note="MIR repeat: matches 1..140 of consensus"
misc_feature complement(39554..39839)
/note="match: GSS A0070531"
repeat_region 41165..41445
/note="L2 repeat: matches 2180..2489 of consensus"
repeat_region 41615..41810
```

```
Query Match 100.0%; Score 21; DB 9; Length 157875;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 ACTTCTTGCGACGCTGTC 21
|||||
Db 143687 ACTTCTTGCGACGCTGTC 143667

RESULT 2
AC020849/c
LOCUS AC020849
DEFINITION Mus musculus clone RP21-43909, WORKING DRAFT SEQUENCE, 55 unordered
pieces.
ACCESSION AC020849
VERSION AC020849.4 GI:9211211
KEYWORDS HTG, HTGS-PHASE1, HTGS-DRAFT.
SOURCE Mus musculus.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 192169)
AUTHORS DOE Joint Genome Institute.
TITLE Unpublished
JOURNAL 2 (bases 1 to 192169)
REFERENCE Direct Submission
AUTHORS DOE Joint Genome Institute.
TITLE Submitted (10-JAN-2000) Production Sequencing Facility, DOE Joint
JOURNAL Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT On Jul 15, 2000 this sequence version replaced gi:9680155.

```
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 1426557
Center clone name: RPCT-21_43909
-----
Summary Statistics
Consensus quality: 148566 bases at least Q40
Consensus quality: 175179 bases at least Q20
Consensus quality: 175179 bases at least Q20
Estimated insert size: 180000; pulse field gel estimation
Estimated insert size: 186769; sum-of-contigs estimation
Quality coverage: 3.49 in Q20 bases; pulse field gel estimation
Quality coverage: 3.37 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 55 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1236: contig of 1236 bp in length
1237
1336: gap of unknown length
1337
2710: contig of 1374 bp in length
2810: gap of unknown length
2811
4014: contig of 1204 bp in length
4015
4114: gap of unknown length
4115
5404: contig of 1290 bp in length
5505
5504: gap of unknown length
7256: contig of 1752 bp in length
7257
7357: gap of unknown length
8829: contig of 1473 bp in length
8830
8930: gap of unknown length
10241: contig of 1312 bp in length
10242
10341: gap of unknown length
12012: contig of 1671 bp in length
12013
12112: gap of unknown length
12113
13606: contig of 1494 bp in length
13607
13706: gap of unknown length
13707
15052: contig of 1346 bp in length
15053
15152: gap of unknown length
15153
16157: contig of 1005 bp in length
16158
16257: gap of unknown length
16258
17580: contig of 1323 bp in length
17581
17680: gap of unknown length
17681
19000: contig of 1320 bp in length
19001
19100: gap of unknown length
19101
20952: contig of 1852 bp in length
20953
21052: gap of unknown length
22866: contig of 1814 bp in length
22867
22956: gap of unknown length
25270: contig of 2304 bp in length
25271
25370: gap of unknown length
25371
26807: contig of 1437 bp in length
26808
26907: gap of unknown length
28347: contig of 1440 bp in length
28348
28447: gap of unknown length
30458: contig of 2011 bp in length
30459
30558: gap of unknown length
31686: contig of 1128 bp in length
31687
31786: gap of unknown length
33238: contig of 1452 bp in length
33239
33338: gap of unknown length
33339
35318: contig of 1980 bp in length
35319
35418: gap of unknown length
35419
36985: contig of 1567 bp in length
36986
37085: gap of unknown length
37086
39161: contig of 2076 bp in length
39162
39261: gap of unknown length
```

GenCore version 5.1.4 p3 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2003, 09:39:57 ; Search time 604.8 Seconds

(without alignments)
866.135 Million cell updates/sec

Title: US-09-935-464-12

Perfect score: 18

Sequence: 1 tggagagcttggggagca 18

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapept 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

GenEmbl:*

1: gb_da:*

2: gb_htg:*

3: gb_in:*

4: gb_cm:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vl:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_din:*

35: em_htg_rtd:*

36: em_htg_man:*

37: em_htg_vrl:*

38: em_sy:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	100.0	157875	9 HS272L16	AL023754 Human DNA
2	18	100.0	215293	2 AC122914	AC122914 Mus muscu
3	17	94.4	78508	2 AC018191	AC018191 Drosophill
4	17	94.4	139924	2 AC116512	AC116512 Mus muscu
5	17	94.4	144723	10 AC005818	AC005818 Mus muscu
6	17	94.4	154639	2 AC090175	AC090175 Homo sapi
7	17	94.4	158458	2 AC007395	AC007395 Homo sapi
8	17	94.4	159836	9 AL392085	AL392085 Human DNA
9	17	94.4	163597	2 AL162589	AL162589 Homo sapi
10	17	94.4	171569	3 AC007811	AC007811 Drosophill
11	17	94.4	172686	2 AC128905	AC128905 Rattus no
12	17	94.4	174215	2 AC094823	AC094823 Rattus no
13	17	94.4	175420	2 AC129161	AC129161 Rattus no
14	17	94.4	176735	3 AC007810	AC007810 Drosophill
15	17	94.4	177096	2 AC113177	AC113177 Mus muscu
16	17	94.4	177461	2 AC099154	AC099154 Rattus no
17	17	94.4	183057	2 AC117152	AC117152 Rattus no
18	17	94.4	186057	2 AC123412	AC123412 Rattus no
19	17	94.4	186284	2 AC094163	AC094163 Rattus no
20	17	94.4	196785	2 AC120065	AC120065 Rattus no
21	17	94.4	197768	2 AL844174	AL844174 Mus muscu
22	17	94.4	204653	10 AC005302	AC005302 Mus muscu
23	17	94.4	210133	9 AC013564	AC013564 Homo sapi
24	17	94.4	218391	3 AE003719	AE003719 Drosophill
25	17	94.4	226919	2 AL845275	AL845275 Mus muscu
26	16.4	91.1	1021	10 AF036008	AF036008 Mus muscu
27	16.4	91.1	5229	9 HSM804796	HSM804796 Homo sapi
28	16.4	91.1	5543	10 AF175432	AF175432 Mus muscu
29	16.4	91.1	5877	10 AF036009	AF036009 Mus muscu
30	16.4	91.1	6000	10 F175410S02	F175410 Mus muscu
31	16.4	91.1	6000	10 AB056445	AB056445 Mus muscu
32	16.4	91.1	21111	9 AL390116	AL390116 Human DNA
33	16.4	91.1	77951	2 AC118467_5	Continuation (6 of
34	16.4	91.1	80146	2 AC106419	AC106419 Rattus no
35	16.4	91.1	84307	2 AC052065	AC052065 Homo sapi
36	16.4	91.1	88557	2 AL354986	AL354986 Homo sapi
37	16.4	91.1	104441	2 AC095721	AC095721 Rattus no
38	16.4	91.1	106639	2 AC121371	AC121371 Rattus no
39	16.4	91.1	107171	2 AP004152	AP004152 Oryza sat
40	16.4	91.1	112392	9 AL337500	AL337500 Human DNA
41	16.4	91.1	128332	2 AC068657	AC068657 Homo sapi
42	16.4	91.1	128517	2 AC094352	AC094352 Rattus no
43	16.4	91.1	131105	9 AC008000	AC008000 Homo sapi
44	16.4	91.1	137849	2 AC105141	AC105141 Homo sapi
45	16.4	91.1	141399	2 AC110221	AC110221 Mus muscu

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
HS272L16	LOCUS	HS272L16	AL023754.1	GI:4007152	HTG; BM600; ca repeat polymorphism; Ca2+/Calmodulin dependent Protein Kinase LIKE gene. Contains ESTs, STSs, GSSs, genomic marker DIS491 and a ca repeat polymorphism, complete sequence.	Human sapiens	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
HS272L16	LOCUS	HS272L16	AL023754.1	GI:4007152	HTG; BM600; ca repeat polymorphism; Ca2+/Calmodulin dependent Protein Kinase LIKE gene. Contains ESTs, STSs, GSSs, genomic marker DIS491 and a ca repeat polymorphism, complete sequence.	Human sapiens	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 157875)
AUTHORS Graham,D.
TITLE Direct Submission
JOURNAL Submitted (27-NOV-1998) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
COMMENT

Requests: clonerequests@sanger.ac.uk
On Dec 12, 1998 this sequence version replaced gi:3873472.

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence is the entire insert of clone 272L16. This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.

272L16 is from the library RPCII constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/VECTOR:PCYPAC2>. This sequence was generated from a human chromosome 1 bacterial clone contig constructed in collaboration by the Sanger Centre chromosome 1 mapping group and Brian Schutte, Bryan Bjork, Kevin Coppage and Jeffrey Murray, Department of Pediatrics, University of Iowa, USA. Further information can be found at <http://www.sanger.ac.uk/HGP/Chrl>.

FEATURES

Source

Location/Qualifiers

1..157875

/organism="Homo sapiens"

/db_xref="taxon:9606"

/map="q32.1-32.3"

/clone="RP1-272L16"

/clone_lib="RPCI-1"

966..1078

/note="LTR repeat: matches 281..410 of consensus"

1082..1506

/note="KSRD repeat: matches 1..426 of consensus"

1585..1875

/note="LTR repeat: matches 1..319 of consensus"

1898..2024

/note="LTR repeat: matches 58..203 of consensus"

2347..2925

/note="MER34 repeat: matches 6..543 of consensus"

3090..3417

/note="LTR repeat: matches 23..374 of consensus"

3677..4074

/note="MER34B repeat: matches 486..902 of consensus"

4089..5338

/note="LTR repeat: matches 358..1643 of consensus"

5339..5532

/note="Alu repeat: matches 101..302 of consensus"

5536..5965

/note="LTR repeat: matches 5..450 of consensus"

5972..6019

/note="LTR repeat: matches 318..365 of consensus"

6020..6459

/note="LTR repeat: matches 1..450 of consensus"

6460..6781

/note="LTR repeat: matches 1..318 of consensus"

6786..6870

/note="MIR repeat: matches 56..142 of consensus"

8020..8057

/note="MIR repeat: matches 95..146 of consensus"

8150..8293

/note="MIR repeat: matches 68..212 of consensus"

8380..9083

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

/note="MER50 repeat: matches 1..711 of consensus"

9084..9303

/note="MIR repeat: matches 13..262 of consensus"

9493..9795

/note="AluX repeat: matches 1..304 of consensus"

10108..10282

/note="L2 repeat: matches 2506..2695 of consensus"

10929..11005

/note="L2 repeat: matches 2671..2748 of consensus"

12092..12129

/note="L9 copies 2 mer ca 84% conserved"

12095..12130

/note="9 copies 4 mer acac 86% conserved"

12722..12855

/note="MER91A repeat: matches 20..186 of consensus"

13108..13221

/note="MIR repeat: matches 80..191 of consensus"

complement(13563..13879)

/note="match: GSSS A0075457 A0076848"

14235..14279

/note="MIR repeat: matches 60..105 of consensus"

14327..14442

/note="MIR repeat: matches 112..249 of consensus"

14495..15046

/note="L2 repeat: matches 2127..2750 of consensus"

16671..16855

/note="MER5A repeat: matches 2..189 of consensus"

17103..17476

/note="L2 repeat: matches 17..422 of consensus"

17679..17938

/note="MIR repeat: matches 7..262 of consensus"

17891..17953

/note="L2 repeat: matches 2694..2749 of consensus"

18650..18758

/note="L2 repeat: matches 2641..2750 of consensus"

19246..19406

/note="MER5A repeat: matches 4..189 of consensus"

19450..19871

/note="L2 repeat: matches 2309..2748 of consensus"

19916..19987

/note="MIR repeat: matches 63..140 of consensus"

20200..20228

/note="MER34 repeat: matches 510..538 of consensus"

20233..20365

/note="MER34 repeat: matches 413..543 of consensus"

20338..20388

/note="LTR29 repeat: matches 454..503 of consensus"

20459..20809

/note="MER47A repeat: matches 2..366 of consensus"

20976..21143

/note="MER34 repeat: matches 6..172 of consensus"

21834..22278

/note="LTR repeat: matches 1..466 of consensus"

22477..22528

/note="26 copies 2 mer ag 79% conserved"

22826..22937

/note="MER81 repeat: matches 1..112 of consensus"

24897..25034

/note="MER3A repeat: matches 30..171 of consensus"

25084..25193

/note="MIR repeat: matches 26..145 of consensus"

25207..25435

/note="MIR repeat: matches 13..259 of consensus"

25919..26363

/note="LTR repeat: matches 68..541 of consensus"

26364..26813

/note="MER42 repeat: matches 1..503 of consensus"

26814..26885

/note="LTR repeat: matches 1..68 of consensus"

26931..27357

/note="L2 repeat: matches 2292..2723 of consensus"

27779..27820

/note="21 copies 2 mer ca 100% conserved"

May 27 11

```

repeat_region 27779..27818
                /note="10 copies 4 mer caca 100% conserved"
repeat_region 29036..29195
                /note="HAL1 repeat: matches 4..163 of consensus"
repeat_region 29263..29500
                /note="HAL1 repeat: matches 141..355 of consensus"
repeat_region 29501..29820
                /note="MUR1A1 repeat: matches 1..365 of consensus"
repeat_region 29821..30076
                /note="HAL1 repeat: matches 355..618 of consensus"
repeat_region 30095..30228
                /note="FLM4 repeat: matches 7..142 of consensus"
repeat_region 30246..30440
                /note="HAL1 repeat: matches 683..874 of consensus"
repeat_region 30539..30788
                /note="Charliella repeat: matches 1189..1455 of consensus"
repeat_region 30789..31099
                /note="ALUSg repeat: matches 1..306 of consensus"
repeat_region 31100..31167
                /note="Charliella repeat: matches 1125..1189 of consensus"
repeat_region 31170..31914
                /note="FLM4 repeat: matches 5448..6190 of consensus"
repeat_region 31923..32141
                /note="FLM2 repeat: matches 4317..4530 of consensus"
repeat_region 32142..32551
                /note="MSRA repeat: matches 5..426 of consensus"
repeat_region 32552..33436
                /note="FLM2 repeat: matches 3451..4317 of consensus"
repeat_region 33437..33756
                /note="ALUy repeat: matches 1..309 of consensus"
repeat_region 33757..36041
                /note="FLM2 repeat: matches 761..3451 of consensus"
repeat_region 37817..38850
                /note="Charliella repeat: matches 1..1142 of consensus"
repeat_region 39293..39437
                /note="MIR repeat: matches 1..140 of consensus"
misc_feature complement(39554..39839)
                /note="match GSS A0070531"
repeat_region 41165..41445
                /note="L2 repeat: matches 2180..2489 of consensus"
repeat_region 41615..41810

```

```

Query Match 100.0%; Score 18; DB 9; length 157875;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 TGGAGCTTGGGGAGCA 18
    |||||
DB 143358 TGGAGCTTGGGGAGCA 143375

```

```

RESULT 2
AC122914 215293 bp DNA linear HTG 28-MAY-2002
LOCUS Mus musculus chromosome UNK clone RP23-26F9, WORKING DRAFT
DEFINITION
SEQUENCE, 11 unordered pieces.
ACCESSION AC122914.1 GI:21218545
VERSION AC122914.1 GI:21218545
KEYWORDS HTG; HTGS-PHASE1; HTGS-DRAFT.
SOURCE Mus musculus.
ORGANISM Mus musculus

```

```

REFERENCE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE The sequence of Mus musculus clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 215293)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (28-MAY-2002) Genome Sequencing Center, 4444 Forest Park
        Parkway, St. Louis, MO 63108, USA
COMMENT

```

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@watson.wustl.edu
----- Project Information -----
Center project name: M_BA0026F09
----- Summary Statistics -----
Sequencing vector: M13; 0%
Sequencing vector: plasmid; 100%
Chemistry: Dye-terminator Big Dye; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 210598 bases at least Q40
Consensus quality: 211324 bases at least Q30
Consensus quality: 211779 bases at least Q20
Insert size: 212000; agarose-fp
Insert size: 214293; sum-of-contigs
Quality coverage: 12.87 in Q20 bases; agarose-fp
Quality coverage: 10.50 in Q20 bases; sum-of-contigs
-----

```

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

```

1 1233: contig of 1233 bp in length
1234 1333: gap of unknown length
1334 2576: contig of 1243 bp in length
2577 4229: contig of 1553 bp in length
4230 4320: gap of unknown length
4330 9108: contig of 4779 bp in length
9109 9208: gap of unknown length
9209 15839: contig of 6631 bp in length
15840 15939: gap of unknown length
15940 26130: contig of 10191 bp in length
26131 26230: gap of unknown length
26231 50043: contig of 23813 bp in length
50044 50143: gap of unknown length
50144 74268: contig of 24125 bp in length
74269 74369: gap of unknown length
74370 101258: contig of 26890 bp in length
101259 101359: gap of unknown length
101360 142855: contig of 41497 bp in length
142856 215293: contig of 72338 bp in length.
Location/Qualifiers
1..215293
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="UNK"
/clone="RP23-26F9"
1..1233
/misc_feature /note="assembly.name:Contig10"
1334..2576
/misc_feature /note="assembly.name:Contig11"
2577..4229
/misc_feature /note="assembly.name:Contig13"
4330..9108
/misc_feature /note="assembly.name:Contig14"
9209..15839
/misc_feature /note="assembly.name:Contig15"
15940..26130
/misc_feature /note="assembly.name:Contig16"
26231..50043
/misc_feature /note="assembly.name:Contig17"
50144..74268
/misc_feature /note="assembly.name:Contig18"

```

http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
 Seq primer: M13 Reverse
 Class: BAC ends.

FEATURES

Location/Qualifiers
 1..388

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="2367C8"
 /clone_lib="CIT-RSP"
 /sex="Male"
 /cell_type="Sperm"
 /note="Vector: pBeloBAC11, Site_1: HindIII, Site_2:
 HindIII"

BASE COUNT 102 a 88 c 105 g 93 t
 ORIGIN

Query Match 100.0%; Score 18; DB 17; Length 388;
 Best Local Similarity 100.0%; Pred. No. 6e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGGGAGCTTGGGGGAGCA 18
 ||||||||||||||||
 Db 271 TGGGAGCTTGGGGGAGCA 288

RESULT 2

A1215131

LOCUS

DEFINITION gp14f02.x1 NCI CGAP C08 Homo sapiens cDNA clone IMAGE:192555 3'
 similar to TR:008763 008763 PROTEIN KINASE ; mRNA sequence.

ACCESSION A1215131

VERSION A1215131.1 GI:3778732
 EST.

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 445)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/BLMT at:
 www-bio.lnll.gov/bdrrp/image/image.html
 Seq primer: -40UP from Gibco
 High quality sequence stop: 444.

FEATURES

Location/Qualifiers

1..445
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:192555"
 /clone_lib="NCI-CGAP_C08"
 /tissue_type="adenocarcinoma"
 /lab_host="DH10B"

/note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a
 modified polylinker; 1st strand cDNA was prepared from
 colon adenocarcinoma, and was then primed with a Not I -
 oligo(dT) primer. Double-stranded cDNA was ligated to Eco
 RI adaptors (Pharmacia), digested with Not I and cloned
 into the Not I and Eco RI sites of the modified pT73
 vector. Library is normalized. Library was constructed by
 Bento Soares and M. Fatima Bonaldo."

BASE COUNT 108 a 91 c 124 g 121 t 1 others

ORIGIN

Query Match 100.0%; Score 18; DB 9; Length 445;
 Best Local Similarity 100.0%; Pred. No. 6.2e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGGGAGCTTGGGGGAGCA 18
 ||||||||||||||||
 Db 233 TGGGAGCTTGGGGGAGCA 250

RESULT 3

A4293753

LOCUS

DEFINITION RPCI-23-148E23.TV RPCI-23 Mus musculus genomic clone RPCI-23-148E23
 , DNA sequence.

ACCESSION A4293753

VERSION A4293753.1 GI:9535627

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Mouse BAC End Sequences from Library RPCI-23
 Unpublished (1999)
 Other GSSs: RPCI-23-148E23.TV
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhaoc@tigr.org
 Clones are derived from the mouse BAC library RPCI-23. For BAC
 library availability, please contact Pieter de Jong
 (pieter@dejong.med.bufileo.edu). Clones may be purchased from
 BACPAC Resources (http://bacpac.med.bufileo.edu/orderingframe.htm)
 or from Resea ch Genetics (info@resgen.com). BAC end page:
 http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
 Plate: 148 row: E column: 23
 Seq primer: SP6
 Class: BAC ends.

FEATURES

Location/Qualifiers

1..480
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="RPCI-23-148E23"
 /clone_lib="RPCI-23"
 /sex="Female"
 /lab_host="DH10B"

/note="Organ: Kidney/Brain; Vector: pBAC3.6; Site_1:
 EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
 brain genomic DNA was isolated and partially digested
 with a combination of EcoRI and EcoRI Methylase. Size
 selected DNA was cloned into the pBAC3.6 vector at the
 EcoRI sites. The ligation products were transformed into
 DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 96 a 141 c 99 g 144 t

ORIGIN

Query Match 94.4%; Score 17; DB 17; Length 480;
 Best Local Similarity 100.0%; Pred. No. 1.7e+03;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGGGAGCTTGGGGGAGC 17
 ||||||||||||||||
 Db 449 TGGGAGCTTGGGGGAGC 465

RESULT 4

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2003, 12:18:47 : Search time 1207.5 Seconds

(without alignments)
281.661 Million cell updates/sec

Title: US-09-935-464-39

Perfect score: 21

Sequence: 1.gataccccgcgtctatgaag 21

Scoring table: IDENTITY_NDC

Gapop 10.0, Gapext 1.0

Searched: 1615406 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

BST:*

- 1: em_estba:*
- 2: em_esthm:*
- 3: em_estln:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_hic:*
- 9: gb_estl:*
- 10: gb_est2:*
- 11: gb_hic:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estlm:*
- 16: em_estom:*
- 17: gb_gss:*
- 18: em_gss_hum:*
- 19: em_gss_juv:*
- 20: em_gss_pla:*
- 21: em_gss_vit:*
- 22: em_gss_fun:*
- 23: em_gss_mam:*
- 24: em_gss_mus:*
- 25: em_gss_other:*
- 26: em_gss_pro:*
- 27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID.	Description
1	21	100.0	388	17	A0077073 CIT-HSP-2
2	21	100.0	573	14	B0086330 B0086330
3	21	100.0	742	13	B1818261 B1818261
4	21	100.0	809	13	B1821474 B1821474
5	19.4	92.4	188	17	A0078706 A0078706
6	19.4	92.4	445	9	A1215131 A1215131

7	19.4	92.4	740	13	B1824483 B1824483
8	19.4	92.4	812	13	B1772626 B1772626
9	19.4	92.4	824	12	B6715920 B6715920
10	19.4	92.4	1126	13	BM547443 BM547443
11	17.8	84.8	521	12	B6738994 B6738994
12	17.8	84.8	627	10	AM826802 AM826802
13	17.8	84.8	637	17	A2569455 A2569455
14	17.8	84.8	657	13	B1960284 B1960284
15	17.8	84.8	699	17	A2573391 A2573391
16	17.4	82.9	823	14	C93862 C93862
17	17.4	81.0	1129	14	BM607335 BM607335
18	16.8	80.0	546	17	AQ242870 AQ242870
19	16.8	80.0	604	17	BH206651 BH206651
20	16.8	80.0	879	14	BQ440480 BQ440480
21	16.8	80.0	920	17	AQ243781 AQ243781
22	16.8	80.0	967	17	CNS030A7 CNS030A7
23	16.4	78.1	235	12	BF176514 BF176514
24	16.4	78.1	718	17	AG182329 AG182329
25	16.4	78.1	801	17	AG876237 AG876237
26	16.4	78.1	1363	17	AG170407 AG170407
27	16.2	77.1	199	9	AA872465 AA872465
28	16.2	77.1	239	14	H34897 H34897
29	16.2	77.1	271	12	BE865002 BE865002
30	16.2	77.1	306	9	A1713924 A1713924
31	16.2	77.1	311	10	BA433820 BA433820
32	16.2	77.1	332	10	AME03037 AME03037
33	16.2	77.1	347	17	CNS070ENT CNS070ENT
34	16.2	77.1	427	17	A0754714 A0754714
35	16.2	77.1	430	17	A2176655 A2176655
36	16.2	77.1	433	17	AQ438334 AQ438334
37	16.2	77.1	437	17	AQ088565 AQ088565
38	16.2	77.1	442	12	BE697081 BE697081
39	16.2	77.1	451	10	AM254051 AM254051
40	16.2	77.1	468	9	AL373446 AL373446
41	16.2	77.1	475	17	FR0049127 FR0049127
42	16.2	77.1	484	17	A2170482 A2170482
43	16.2	77.1	499	17	A2386531 A2386531
44	16.2	77.1	526	13	BI449923 BI449923
45	16.2	77.1	552	12	B6371336 B6371336

ALIGNMENTS

RESULT 1
A0077073
LOCUS 388 bp DNA linear GSS 20-AUG-1998
DEFINITION CIT-HSP-2367C8.TR CIT-HSP Homo sapiens genomic clone 2367C8. DNA
ACCESSION A0077073 GI:3438257
VERSION A0077073.1
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 388)
AUTHORS Adams M.D., Rounsley S.D., Zhao S., Bass S., Liber K., Golden K., Berry K., Granger D., Suh E., Wible C., Shizuya H., Simon M. and Venter J.C.
TITLE use of a random human BAC End Sequence Database for Sequence-Ready Map Building
JOURNAL Unpublished (1998)
COMMENT Other GSSs: CIT-HSP-2367C8.TF
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@igf.org
Clones are available from Research Genetics (info@resgen.com). BAC end search page:

http://www.flycr.org/cdb/hungen/bac_end_search/bac_end_search.html.

Seq primer: M13 Reverse

Class: BAC ends.

FEATURES

Source

Location/Qualifiers
1..388

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="CIT-HSP"

/sex="Male"

/cell_type="Sperm"

/note="Vector: pBelobAC11, Site_1: HindIII, Site_2:
HindIII"

BASE COUNT 102 a 88 c 105 g 93 t

ORIGIN

Query Match 100.0%; Score 21; DB 17; Length 388;
Best Local Similarity 100.0%; Pred. No. 5.1;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATACCCCGCTCTATGAAG 21

Db 360 GATACCCCGCTCTATGAAG 380

RESULT 2

LOCUS

B0086330 573 bp mRNA linear EST 29-APR-2002

DEFINITION 1j2c07.y1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens

CDNA clone IMAGE:615181 5' similar to TR:095523 095523 D0272L16.1

EST.

ACCESSION B0086330

VERSION B0086330.1

KEYWORDS GI:20045534

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 573)

Melton,D., Brown,J., Kenty,G., Permut,A., Lee,C., Kaestner,K.,

Lemishka,I., Scaerce,M., Brestelli,J., Gradwohl,G., Clifton,S.,

Hillier,L., Maira,M., Pape,D., Wylie,T., Martin,J., Blistein,A.,

Schmitt,A., Pfeising,B., Ritter,E., Ronno,I., Bennett,J., Cardenas

,M., Gibbons,M., McCann,R., Cole,R., Tsagarelisvilli,R., Williams,T.

, Jackson,K. and Bowers,Y.

Endocrine Pancreas Consortium

Unpublished (2000)

CONTACT: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Endocrine Pancreas Consortium

Harvard University, Howard Hughes Medical Institute

Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,

MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@bioh.harvard.edu

Library was constructed by Dr. Douglas Melton DNA sequencing by:

Washington University Genome Sequencing Center This clone is

available royalty-free through LNL; please contact the IMAGE

Consortium (info@image.lnl.gov) for further information

Seq primer: -408P from Gibco

High quality sequence stop: 447.

Location/Qualifiers

1..573

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:6135181"

/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"

/sex="Both"

/tissue_type="Islets of Langerhans"

/dev_stage="Adult"

/lab_host="DH10B"

/note="Organ: Pancreas; Vector: pSPORT1; Site_1: Not 1;
Site_2: Sal 1; Starting library constructed using

SuperScript Plasmid Library kit (Life Technologies). cDNA

made by oligo-dT priming. Size selected by column

fractionation; average insert size 1.08 kb. Library was

amplified once on solid support and plasmid DNA from

library was prepared. The library DNA was normalized by

method #4 from Bonaldo, Lennon, and Soares 1996 Genome

Research 6:791-806; 0.5 microgram single-stranded library

plasmid DNA was mixed with 5 micrograms PCR product

representing library inserts and hybridized to an Ecot of

20. Single-stranded (unhybridized) plasmids were isolated

by hydroxyapatite chromatography and used to make this

library."

BASE COUNT 143 a 177 c 136 g 117 t

ORIGIN

Query Match 100.0%; Score 21; DB 14; Length 573;
Best Local Similarity 100.0%; Pred. No. 5.6;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATACCCCGCTCTATGAAG 21

Db 149 GATACCCCGCTCTATGAAG 169

RESULT 3 742 bp mRNA linear EST 04-OCT-2001

LOCUS B1818261

DEFINITION 603032510F1 NIH_MGC_115 Homo sapiens CDNA clone IMAGE:5173587 5',

mRNA sequence.

ACCESSION B1818261

VERSION B1818261.1

KEYWORDS GI:15928724

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 742)

NIH-MGC http://mgc.nci.nih.gov/

CONTACT: Robert Strausberg, Ph.D.

Email: c9apbs-remail.nih.gov

Tissue Procurement: Life Technologies, Inc.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Inocyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

http://image.lnl.gov

Plate: L14M11432 row: d column: 04

High quality sequence stop: 742.

Location/Qualifiers

1..742

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5173587"

/clone_lib="NIH_MGC_115"

/lab_host="DH10B"

/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site_1: Not; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally inserted (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb.
Insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

BASE COUNT 197 a 180 c 197 g 168 t

ORIGIN

Query Match 100.0%; Score 21; DB 13; Length 742;
Best Local Similarity 100.0%; Pred. No. 6;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATACCCCGCTCTATGAAG 21

Db 149 GATACCCCGCTCTATGAAG 169

GenCore version 5.1.4.P5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2003, 12:18:47 ; Search time 1035 Seconds

(without alignments)
281.661 Million cell updates/sec

Title: US-09-935-464-12

Sequence: 1 tgggagcttggggagca 18

Scoring table: IDENTITY_NUC

Gapop 10.0 ; Gapext 1.0

Searched: 1615406 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estda:*
2: em_esthum:*
3: em_estlin:*
4: em_estma:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pin:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mem:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	18	100.0	388	17	A0077073
2	18	100.0	445	9	AT215131
3	17	94.4	480	17	A2293753 RPCI-23-1
4	17	94.4	702	17	A2858794 2M0164D11
5	17	94.4	706	17	A2117160 RPCI-23-4
6	17	94.4	919	12	BG259997 602371772

7	16.4	91.1	105	10	AM566487
8	16.4	91.1	326	14	H62161
9	16.4	91.1	421	12	BF872369
10	16.4	91.1	437	14	H54659
11	16.4	91.1	468	17	A2791804
12	16.4	91.1	495	17	AQ140076
13	16.4	91.1	502	14	H62160
14	16.4	91.1	508	17	BH282230
15	16.4	91.1	520	14	T98678
16	16.4	91.1	547	17	A2791792
17	16.4	91.1	603	14	W27259
18	16.4	91.1	658	17	AG160341
19	16.4	91.1	704	9	A1140100
20	16.4	91.1	707	10	BE296441
21	16.4	91.1	714	10	BE269258
22	16.4	91.1	740	17	BH056996
23	16.4	91.1	771	17	AG688518
24	16.4	91.1	772	17	AG180243
25	16.4	91.1	817	10	BE407579
26	16.4	91.1	837	12	BG164441
27	16.4	91.1	904	12	BG110945
28	16.4	91.1	914	14	BQ952330
29	16.4	91.1	1032	12	BG766985
30	16.4	91.1	1046	13	BM470117
31	16.4	91.1	1194	13	BM548548
32	16.4	91.1	1244	12	BF921580
33	16.4	91.1	1255	10	BE202729
34	16.4	91.1	1263	10	AM437750
35	16.4	91.1	1284	17	A2400996
36	16.4	91.1	1287	14	BM936237
37	16.4	91.1	1300	9	AU231985
38	16.4	91.1	1333	10	BE233412
39	16.4	91.1	1350	9	AU234045
40	16.4	91.1	1350	9	AM227235
41	16.4	91.1	1358	9	AU278446
42	16.4	91.1	1360	13	BM429447
43	16.4	91.1	1394	13	BI540154
44	16.4	91.1	1411	13	BI404426
45	16.4	91.1	1412	13	BI540155

ALIGNMENTS

RESULT 1
A0077073
LOCUS
DEFINITION
CMT-HSE-2367C8.TR CTF-HSP Homo sapiens genomic clone 2367C8, DNA
sequence.
ACCESSION
A0077073
VERSION
A0077073.1 GI:3438257
KEYWORDS
GSS.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 388)
ADAMS M.D., Rounsley S.D., Zhao S., Bass S., Linher K., Golden K.,
Berry K., Granger D.R., Suh E., Whible C., Shizuya H., Simon M. and
Venter J.C.
Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
Unpublished (1998)
JOURNAL
Other GSSs: CTF-HSP-2367C8.TF
COMMENT
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: m.adams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:

http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
 Seq primer: M13 Reverse
 Class: BAC ends.

FEATURES
 Location/Qualifiers
 1.388

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="2367c8"
 /clone_lib="CTT-HSP"
 /sex="Male"
 /cell_type="Sperm"
 /note="Vector: pBelbac11; site_1: HindIII; site_2: HindIII"

BASE COUNT 102 a 88 c 105 g 93 t
 ORIGIN

Query Match 100.0%; Score 18; DB 17; Length 388;
 Best Local Similarity 100.0%; Pred. No. 6e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGGAGCTTGGGGAGCA 18
 Db 271 TGGAGCTTGGGGAGCA 288

RESULT 2 445 bp mRNA linear EST 21-OCT-1998
 AI215131
 LOCUS gp41f02.x1 NCI-CGAP_C08 Homo sapiens CDNA clone IMAGE:192555 3'
 DEFINITION similar to TR:008763 008763 PROTEIN KINASE ; , mRNA sequence.
 ACCESSION AI215131
 VERSION AI215131
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 445)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.
 DNA Sequencing by: Greg Lennon, Ph.D.
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINL at:
 www.bio.lnlnl.gov/bdrp/image/image.html
 Seq primer: 400P from Gibco
 High quality sequence stop: 444.

FEATURES
 SOURCE

Location/Qualifiers
 1.445

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="NCI-CGAP_C08"
 /clone_lib="NCI-CGAP_C08"
 /tissue_type="adenocarcinoma"
 /lab_host="DH10B"

/note="Organ: colon; Vector: p773D-Pac (Pharmacia) with a
 modified polylinker; 1st strand cDNA was prepared from
 colon adenocarcinoma, and was then primed with a Not I -
 oligo(dT) primer. Double-stranded cDNA was ligated to Eco
 RI adaptors (Pharmacia), digested with Not I and cloned
 into the Not I and Eco RI sites of the modified p773
 vector. Library is normalized. Library was constructed by
 Bento Soares and M. Fatima Bernaldo."
 BASE COUNT 108 a 91 c 124 g 121 t 1 others
 ORIGIN

Query Match 100.0%; Score 18; DB 9; Length 445;
 Best Local Similarity 100.0%; Pred. No. 6.2e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGGAGCTTGGGGAGCA 18
 Db 233 TGGAGCTTGGGGAGCA 250

RESULT 3 480 bp DNA linear GSS 27-JUL-2000
 A2293753
 LOCUS RPECI-23-148E23.TV RPECI-23 Mus musculus genomic clone RPECI-23-148E23
 DEFINITION RPECI-23-148E23.TV RPECI-23 Mus musculus genomic clone RPECI-23-148E23
 , DNA sequence.
 ACCESSION A2293753
 VERSION A2293753
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 480)
 Zhao, S., Nieman, W., Feldblyum, T., Malek, J., Shatsman, S., Aknret,
 B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Kroi, M., de Jong, P.
 and Fraser, C.M.
 Mouse BAC End Sequences from Library RPECI-23
 Unpublished (1999)
 Other-GSS: RPECI-23-148E23.TV
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: zshaof@tigr.org
 Clones are derived from the mouse BAC library RPECI-23. For BAC
 library availability, please contact Pieter de Jong
 (pieter@edj.med.buffalo.edu). Clones may be purchased from
 BACPAC Resources (http://bacpac.med.buffalo.edu/ordering.htm)
 or from Resea ch Genetics (info@resgen.com). BAC end page:
 http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
 Plate: 148 row: E column: 23
 Seq primer: SP6
 Class: BAC ends.

FEATURES
 SOURCE

Location/Qualifiers
 1.480

/organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="RPECI-23-148E23"
 /clone_lib="RPECI-23"
 /sex="Female"
 /lab_host="DH10B"
 /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site 1:
 EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or
 brain genomic DNA was isolated and partially digested
 with a combination of EcoRI and EcoRI Methylase. Size
 selected DNA was cloned into the pBACe3.6 vector at the
 EcoRI sites. The ligation products were transformed into
 DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 96 a 141 c 99 g 144 t
 ORIGIN

Query Match 94.4%; Score 17; DB 17; Length 480;
 Best Local Similarity 100.0%; Pred. No. 1.7e+03;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGGAGCTTGGGGAGC 17
 Db 449 TGGAGCTTGGGGAGC 465

RESULT 4

May

```

repeat_region 2779..27818
/note="10 copies 4 mer caca 100% conserved"
repeat_region 29036..29195
/note="HAL1 repeat: matches 4..163 of consensus"
repeat_region 29263..29500
/note="HAL1 repeat: matches 141..355 of consensus"
repeat_region 29501..29820
/note="HAL1 repeat: matches 1..365 of consensus"
repeat_region 29821..30076
/note="HAL1 repeat: matches 355..618 of consensus"
repeat_region 30095..30228
/note="HAL1 repeat: matches 7..142 of consensus"
repeat_region 30246..30440
/note="HAL1 repeat: matches 683..874 of consensus"
repeat_region 30539..30788
/note="HAL1 repeat: matches 1189..1455 of consensus"
repeat_region 30789..31099
/note="HAL1 repeat: matches 1..306 of consensus"
repeat_region 31100..31167
/note="HAL1 repeat: matches 1125..1189 of consensus"
repeat_region 31170..31914
/note="HAL1 repeat: matches 5448..6190 of consensus"
repeat_region 31923..32141
/note="HAL1 repeat: matches 4317..4530 of consensus"
repeat_region 32142..32551
/note="HAL1 repeat: matches 5..426 of consensus"
repeat_region 32552..33436
/note="HAL1 repeat: matches 3451..4317 of consensus"
repeat_region 33437..33756
/note="HAL1 repeat: matches 1..309 of consensus"
repeat_region 33757..33804
/note="HAL1 repeat: matches 761..3451 of consensus"
repeat_region 37817..38850
/note="HAL1 repeat: matches 1..1142 of consensus"
repeat_region 39293..39437
/note="HAL1 repeat: matches 1..140 of consensus"
misc_feature complement(39534..39839)
/note="match: GSS A0070531"
repeat_region 41165..41445
/note="L2 repeat: matches 2180..2489 of consensus"
repeat_region 41615..41810
/note="L2 repeat: matches 2180..2489 of consensus"

Query Match 100.0%; Score 18; DB 9; Length 157875;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGGGACCTTGGGGAGCA 18
Db 143358 TGGGACCTTGGGGAGCA 143375

RESULT 2
AC122914 215293 bp DNA linear HMG 28-MAY-2002
LOCUS Mus musculus chromosome UNK clone RP23-26F9, WORKING DRAFT
DEFINITION AC122914.1 GI:21218545
VERSION AC122914.1 GI:21218545
KEYWORDS HMG; HMGs; PHASE1; HMGs; DRAFT.
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 215293)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE The sequence of Mus musculus clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 215293)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (28-MAY-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
COMMENT

```

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.wustl.edu
----- Project Information -----
Center project name: M_BA0026F09
----- Summary Statistics -----
Sequencing vector: M13, 0%
Sequencing vector: plasmid, 100%
Chemistry: Dye-terminator ET, 0% of reads
Chemistry: Dye-terminator Big Dye, 100% of reads
Assembly program: Phrap, version 0.990319
Consensus quality: 210596 bases at least Q40
Consensus quality: 211324 bases at least Q30
Consensus quality: 211779 bases at least Q20
Insert size: 212000; agarose-1p
Insert size: 214293; sum-of-contigs
Quality coverage: 12.87 in Q20 bases; agarose-1p
Quality coverage: 10.50 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
-----
1 1233: contig of 1233 bp in length
1 1234: gap of unknown length
1 1334: contig of 1243 bp in length
1 2576: gap of unknown length
1 2577: gap of unknown length
1 2677: contig of 1553 bp in length
1 4329: gap of unknown length
1 4330: gap of unknown length
1 9109: contig of 4779 bp in length
1 9209: gap of unknown length
1 9209: gap of unknown length
1 15840: contig of 6631 bp in length
1 15939: gap of unknown length
1 26130: contig of 10191 bp in length
1 26230: gap of unknown length
1 26231: gap of unknown length
1 50043: contig of 23813 bp in length
1 50143: gap of unknown length
1 50144: gap of unknown length
1 74269: contig of 24125 bp in length
1 74369: gap of unknown length
1 74369: gap of unknown length
1 101259: contig of 26890 bp in length
1 101259: gap of unknown length
1 101359: gap of unknown length
1 142856: contig of 41497 bp in length
1 142856: gap of unknown length
1 142956: 215293: contig of 72338 bp in length.
-----
FEATURES
source
1..215293
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="UNK"
/clone="RP23-26F9"
1..1233
/note="assembly_name:Contig10"
1334..2576
/note="assembly_name:Contig11"
2577..4229
/note="assembly_name:Contig13"
4330..9108
/note="assembly_name:Contig14"
9209..15839
/note="assembly_name:Contig15"
15940..26130
/note="assembly_name:Contig16"
26231..50043
/note="assembly_name:Contig17"
50144..74268
/note="assembly_name:Contig18

```

```

misc_feature       clone_end:T7
                    vector_side:left"
74369..101258
misc_feature       /note="assembly_name:Contig19"
101359..142855
misc_feature       /note="assembly_name:Contig20"
142956..215293
misc_feature       /note="assembly_name:Contig21"
62075 a 44444 c 42411 g 65360 t 1003 others
BASE COUNT
Query Match      100.0%; Score 18; DB 2; Length 215293;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGGGAGCTTGGGGAGCA 18
|||||
Db 44171 TGGGAGCTTGGGGAGCA 44188

RESULT 3
AC018191/c
LOCUS
DEFINITION
AC018191
Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered
pieces.
AC018191
AC018191.1 GI:6553000
HTG: HTGS_PHASE2.
Drosophila melanogaster.
SOURCE
ORGANISM
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
AUTHORS
1 (bases 1 to 78508)
TITLE
Direct Submission
JOURNAL
Submitted (09-DEC-1999) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
COMMENT
This sequence was identified as CDY:10213727 by the submitter.
For more information on this record e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
Source
Location/Qualifiers
1..78508
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
BASE COUNT
23544 a 16827 c 16845 g 21292 t
ORIGIN
Query Match      94.4%; Score 17; DB 2; Length 78508;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GGGAGCTTGGGGAGCA 18
|||||
Db 72959 GGGAGCTTGGGGAGCA 72943

RESULT 4
AC116512/c
LOCUS
DEFINITION
AC116512
Mus musculus clone RP24-328C19, WORKING DRAFT SEQUENCE, 9 ordered
pieces.
AC116512
AC116512.3 GI:21735332
HTG: HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 139924)

```

AUTHORS TITLE JOURNAL REFERENCE AUTHORS

Birren, B., Nusbaum, C. and Lander, E.
Mus musculus, clone RP24-328C19
Unpublished
2 (bases 1 to 139924)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
Boukhalter, B., Brown, A., Camarato, J., Campolano, A., Chang, J.,
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collamore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Fero, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Grinde, S., Goid, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Larocque, K., Lamazares, P.,
Landers, T., Lechoczky, J., Levine, R., Lindblad-Toh, K., Liu, G.,
Maclean, C., MacDonald, P., Major, J., Margulis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., Melidim, J., Meneus, L.,
Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Sudirmanian, A., Talamas, J., Testaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Triggilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE JOURNAL AUTHORS

Submitted (28-MAR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 139924)

Barn, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B.,
Camarato, J., Chang, J., Chazaro, B., Choepel, Y., Collamore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Fero, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J.,
Gardyna, S., Goid, S., Graham, L., Grand-Pierre, N., Hagos, B.,
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,
Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
Liu, G., Maclean, C., MacDonald, P., Major, J., Matthews, C.,
McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mienga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Roy, A., Schauer, S., Schnupack, R., Seaman, S., Severy, P.,
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.

TITLE JOURNAL

COMMENT

Submitted (12-JUL-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 12, 2002 this sequence version replaced g1:21700663.
All repeats were identified using RepeatMasker:
Smith, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: 125512
Center clone name: 328.C.19
Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 137116 bases at least Q40
Consensus quality: 138355 bases at least Q30
Consensus quality: 138779 bases at least Q20
Insert size: 132000; agarose-tp
Insert size: 139124; sum-of-ctnigs
Quality coverage: 8.1 in Q20 bases; agarose-tp

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2003, 09:39:57 ; Search time 705.6 seconds

(without alignments)
866.155 Million cell updates/sec

Title: US-09-935-464-39

Perfect score: 21

Sequence: 1 gatccccccgttattgaag 21

Scoring table: IDENTITY_NUC

Gapop 10.0 ; Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

GenBank:*

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_cm:*

5: gb_ov:*

6: gb_pal:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_dat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vi:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pla:*

35: em_htg_rdg:*

36: em_htg_mam:*

37: em_htg_vtl:*

38: em_sy:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	2447	6	AX399682
2	21	100.0	2464	9	AF428261
3	21	100.0	2474	9	BC032787
4	19.4	92.4	1738	9	HS272161
5	19.4	92.4	2612	9	AK095713
6	19.4	92.4	157875	9	HS272161
7	17.4	82.9	52478	10	AL672172
8	17.4	82.9	167353	2	AC101972
9	17.4	82.9	172273	2	AL845441
10	17.4	82.9	216502	2	AC025964
11	17.4	82.9	240536	10	AC025910
12	17	81.0	459	6	AX433172
13	16.8	80.0	80049	2	AC103139
14	16.8	80.0	109882	2	AC111359
15	16.8	80.0	161529	2	AC096081
16	16.8	80.0	177639	2	AC130761
17	16.8	80.0	337261	2	AC096079
18	16.2	77.1	199	11	G73213
19	16.2	77.1	947	9	HS4323617
20	16.2	77.1	1013	10	D86557
21	16.2	77.1	1032	10	AB023027
22	16.2	77.1	1332	10	D86556
23	16.2	77.1	1402	10	RATPRK1
24	16.2	77.1	1439	10	RATPRK1
25	16.2	77.1	1554	10	AF181984
26	16.2	77.1	1618	10	AB004267
27	16.2	77.1	2333	3	LMA297564
28	16.2	77.1	2416	10	BC021840
29	16.2	77.1	2427	10	AF428262
30	16.2	77.1	3811	3	LM091743
31	16.2	77.1	36068	3	AC012051
32	16.2	77.1	37778	3	AC009602
33	16.2	77.1	42070	8	SPCC1322
34	16.2	77.1	89354	9	AC008957
35	16.2	77.1	115626	9	AC010651
36	16.2	77.1	135003	2	AC112475
37	16.2	77.1	153452	2	AC044895
38	16.2	77.1	154334	2	AC122530
39	16.2	77.1	155937	2	AC016341
40	16.2	77.1	156940	9	AC007444
41	16.2	77.1	160025	2	AC120619
42	16.2	77.1	163046	9	AP000889
43	16.2	77.1	163054	2	AC128087
44	16.2	77.1	165229	2	AC095327
45	16.2	77.1	166668	2	AC126913

ALIGNMENTS

RESULT 1	AX399682	2447 bp	DNA	linear	PAT 06-JUN-2002
LOCUS	AX399682				
DEFINITION	Sequence 3 from Patent WO0224947.				
ACCESSION	AX399682				
VERSION	AX399682.1	GI:21335455			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
TITLE	Delaney, A.D. and Yoganathan, T.				
JOURNAL	Cancer associated protein kinases and their uses				
	Patent: WO 0224947-A 3 28-MAR-2002;				

FEATURES KINETEK PHARMACEUTICALS INC (CA); UNIV BRITISH COLUMBIA (CA)
 Location/Qualifiers
 1..2447
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 70..1500
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD33521.1"
 /db_xref="GI:21335456"
 /translation="MGREEDDCSSMKKOTNIRKPTFMEVLGSGAFSEVPLVQRL
 TGLFALCKIKSPAFRDSLENIIVALKIKHENIVTLIEDIYESTTHYLVQLVSG
 GELFDRLERGVTEKDSAVIOOVLSAKYIHENGIVHDLKPEMLILYTPENSKI
 MTDRLSLMEQNGMNSTAGTPTGYAPVLAQKPYSAVDCWSIGVITVILCGYPP
 FYEETSKLFEKIKGPIYERESPMDISESKDPICHLEKDPNERNYCEKALSHPW
 IDENLALHDIYPSVSLQIQNFKAKKMQANAAVYHHRKILHNLHSPGPIREVE
 NRPEIQASETSRSPSEITITTEAPVLDHVALPVLQLPCHGRRLTAPGGRSLNCL
 VNSLHSSSLVPMHQSILAGCGCCSCLINIGKSKSSYCSPEPLTKRANKKONFK
 SEVMPVYKASGSSHCRAGQGVCLIM"

BASE COUNT 590 a 707 c 604 g 546 t
 ORIGIN

Query Match 100.0%; Score 21; DB 6; Length 2447;
 Best Local Similarity 100.0%; Pred. No. 0.84;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATACCCCGCTCTATGAG 21
 Db 713 GATACCCCGCTCTATGAG 733

RESULT 2
 LOCUS AF428261 2464 bp mRNA linear PRI 07-NOV-2001
 DEFINITION Homo sapiens calcium/calmodulin-dependent protein kinase I gamma
 (CAMK1G) mRNA, complete cds.
 ACCESSION AF428261
 VERSION AF428261.1 GI:16755791
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 2464)
 Schutte,B.C., Bjork,B.C., Coppage,K.B., Malik,M.I., Gregoroy,S.G.,
 Scott,D.J., Brentzell,L.M., Watanabe,Y., Dixon,M.J., and Murray,J.C.
 A preliminary gene map for the Van der Woude syndrome critical
 region derived from 900 kb of genomic sequence at 1q32-q41
 Genome Res. 10 (1), 81-94 (2000)

JOURNAL MEDLINE 20113118
 PUBMED 10645953
 REFERENCE 2 (bases 1 to 2464)
 Bjork,B.C., Watanabe,Y., Murray,J.C. and Schutte,B.C.
 Characterization of the human ortholog of rat Cam Kinase I gamma
 (CamK1g) at 1q32-q41
 Unpublished
 3 (bases 1 to 2464)
 Bjork,B.C., Watanabe,Y., Murray,J.C. and Schutte,B.C.
 Direct Submission
 TITLE Submitted (05-OCT-2001) Genetics, University of Iowa, 140 ENRB,
 Iowa City, IA 52242, USA
 JOURNAL FEATURES
 source
 1..2464
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /map="1q32-q41"
 1..2464
 /gene="CAMK1G"
 73..1503
 /gene="CAMK1G"
 /codon_start=1
 /product="Calcium/calmodulin-dependent protein kinase I"

gamma"
 /protein_id="AAL28100.1"
 /db_xref="GI:16755792"
 /translation="MGREEDDCSSMKKOTNIRKPTFMEVLGSGAFSEVPLVQRL
 TGLFALCKIKSPAFRDSLENIIVALKIKHENIVTLIEDIYESTTHYLVQLVSG
 GELFDRLERGVTEKDSAVIOOVLSAKYIHENGIVHDLKPEMLILYTPENSKI
 MTDRLSLMEQNGMNSTAGTPTGYAPVLAQKPYSAVDCWSIGVITVILCGYPP
 FYEETSKLFEKIKGPIYERESPMDISESKDPICHLEKDPNERNYCEKALSHPW
 IDENLALHDIYPSVSLQIQNFKAKKMQANAAVYHHRKILHNLHSPGPIREVE
 NRPEIQASETSRSPSEITITTEAPVLDHVALPVLQLPCHGRRLTAPGGRSLNCL
 VNSLHSSSLVPMHQSILAGCGCCSCLINIGKSKSSYCSPEPLTKRANKKONFK
 SEVMPVYKASGSSHCRAGQGVCLIM"
 2415..2420
 /gene="CAMK1G"

polyA_signal
 BASE COUNT 609 a 708 c 602 g 545 t
 ORIGIN

Query Match 100.0%; Score 21; DB 9; Length 2464;
 Best Local Similarity 100.0%; Pred. No. 0.84;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATACCCCGCTCTATGAG 21
 Db 716 GATACCCCGCTCTATGAG 736

RESULT 3
 LOCUS BC032787 2474 bp mRNA linear PRI 27-JUN-2002
 DEFINITION Homo sapiens, calcium/calmodulin-dependent protein kinase IG, clone
 MGC:44894 IMAGE:5179957, mRNA, complete cds.
 ACCESSION BC032787
 VERSION BC032787.1 GI:21619664
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 2474)
 Strausberg,R.
 Direct Submission
 Submitted (06-JUN-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
 Sequencing Center (NISC),
 Gaithersburg, Maryland.
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgcgen@lml.nih.gov
 Ahlert,N., Ayala,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
 Blakesley,R.W., Bouffard,G.G., Breun,K., Brinkley,C., Brooks,S.,
 Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
 Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Latic,P., Legaspi,R.,
 Maduro,O.L., Mastello,C., Maskeri,B., Mastlian,S.D., McCloskey,J.C.,
 McDowell,J., Pearson,R., Stantrop,S., Thomas,P.J., Touchman,J.W.,
 Tsirepan,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
 Young,A., Zhang,L.-H. and Green,E.D.

REMARK
 COMMENT
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRK Plate: 68 Row: K Column: 17
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 14196444.
 1..2474
 /organism="Homo sapiens"

	10	16.2	77.1	158	22	AAtK41403	Human bone marrow
	11	16.2	77.1	158	22	AAI22149	Probe #12083 used
	12	16.2	77.1	158	22	AAI67447	Probe #16133 used
	13	16.2	77.1	158	22	AAI67850	Probe #7841 used
	14	16.2	77.1	158	22	ABSI15409	Human genome-der
	15	16.2	77.1	396	22	ABA44162	Human breast cell
	16	16.2	77.1	396	22	ABA54611	Human foetal liver
	17	16.2	77.1	396	22	ABA24396	Probe #2862 for c
	18	16.2	77.1	396	22	AAK2992	Human brain expr
	19	16.2	77.1	396	22	AAK28344	Human bone marrow
	20	16.2	77.1	396	22	AAI12910	Probe #2843 for c
	21	16.2	77.1	396	22	AAI34270	Probe #2956 used
	22	16.2	77.1	396	22	AAI028829	Probe #2820 used
	23	16.2	77.1	396	22	ABS02853	Human genome-der
	24	16.2	77.1	1458	19	AAV18867	Human
	25	16.2	77.1	9203	22	AAK87134	Calmduilin-depen
	26	15.8	75.2	1862	22	ABK34646	Human Immune/hae
C	27	15.4	73.3	160	20	AAHB6170	Human cpna for n
	28	15.4	73.3	561	24	ABK39303	Human single nuc
	29	15.4	73.3	599	22	AAK8744	DNA encoding lun
	30	15.4	73.3	599	22	AAK81778	Human digestive ;
	31	15.4	73.3	599	24	ABN80133	Human Liver assoc
	32	15.4	73.3	765	20	AAZ16070	Human Liver anti
	33	15.4	73.3	765	20	AAK88684	Human gene expres
	34	15.4	73.3	769	20	AAZ16307	Human validated t
	35	15.4	73.3	773	20	AAZ15722	Human gene expres
	36	15.4	73.3	773	20	AAK98665	Human gene expres
C	37	15.4	73.3	812	20	AAK98089	Human validated t
	38	15.4	73.3	836	20	AAK99082	Human cancer cell
	39	15.4	73.3	839	20	AAZ17738	Human cancer cell
	40	15.4	73.3	880	20	AAZ17208	Human gene expres
	41	15.4	73.3	880	20	AAZ17209	Human gene expres
	42	15.4	73.3	880	20	AAZ17210	Human gene expres
	43	15.4	73.3	1773	22	AAK94826	Human full-lenpti
	44	15.4	73.3	4450	22	AAK90986	Human digestive ;
	45	15.4	73.3	4450	22	AAK92021	Human liver asso

ALIGNMENTS

RESULT 1	
ID	AAD36140 standard; DNB; 2447 BP.
XX	AAD36140
AC	AAD36140;
DT	09-AUG-2002 (first entry)
DE	Human calmodulin kinase, CAMK-XI gene.
KW	Human; cytosolic; antisense gene therapy; screening; protein kinase;
KW	cancer; liver; colon; tumour; inflammation; arthritic synovium; CAMK-XI,
KW	calmodulin kinase; enzyme; gene; chromosome 1q32.1-32.3; ds.
OS	Homo sapiens.
XX	
XX	
FH	Key
FT	Location/Qualifiers
FT	CDS 70..1500
FT	/tag= a
FT	/product= "Human CAMK-XI protein"
PN	WO200224947-A2.
PD	
FD	28-MAR-2002.
PF	20-SEP-2001; 2001WO-IB02237.
XX	
PR	20-SEP-2000; 2000US-23399P.
PR	02-OCT-2000; 2000US-237419P.
PR	02-OCT-2000; 2000US-237423P.
PR	04-OCT-2000; 2000US-238558P.
PR	10-MAY-2001; 2001US-290555P.

XX (KINE-) KINETER PHARM INC.
 PA (CYBR-) UNIV BRITISH COLUMBIA.
 XX Yoganathan T, Delaney AD;
 XX WPI: 2002-394145/42.
 DR P-PSDB: AAE22764.
 XX
 PT Diagnosing cancer, comprises determining the upregulation of expression
 PT of a nucleic acid sequence encoding a protein kinase or upregulation of
 PT expression of the protein kinase, in the cancer
 XX
 PS Claim 16; Page 62-64; 87pp; English.
 XX
 CC The invention relates to a method for screening biologically active agent
 CC that modulates cancer associated protein kinase function. The invention
 CC also relates to a method for diagnosing cancer comprising determining the
 CC upregulation of expression of a nucleic acid sequence encoding a protein
 CC kinase. The method is useful for diagnosing cancer. A protein kinase is
 CC useful for screening biological agents that modulate cancer associated
 CC protein kinase function. Downregulating the activity of protein kinase is
 CC useful for inhibiting the growth of a cancer cell, e.g. liver or colon
 CC cancer. A nucleic acid encoding protein kinase is useful to screen biopsy
 CC derived tumours and inflammatory samples such as arthritic synovium, for
 CC amplified DNA in the cell or increased expression of corresponding mRNA
 CC or protein and is also useful to detect differences in expression levels
 CC such as molecular weight, amino acid and nucleotide sequences between the
 CC two cells. The present sequence is human calmodulin kinase CAMK-XI gene
 CC located on chromosome 1q21.1-32.3.
 CC
 SQ Sequence 2447 BP; 590 A; 707 C; 604 G; 546 T; 0 other;
 Query Match 100.0%; Score 21; DB 24; Length 2447;
 Best Local Similarity 100.0%; Pred. No. 0.22;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATACCCCCCGTCTATGAG 21
 Db 713 GATACCCCCCGTCTATGAG 733

RESULT 2
 AAI60703
 ID AAI60703 standard; cDNA; 1956 BP.
 XX
 AC AAI60703;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 4692.
 XX
 KW Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US34263.
 XX
 PR 21-JAN-2000; 2000US-0489725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX
 XX (HXSE-) HXSEQ INC.
 XX
 PI Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Dzmanac RT;
 XX
 DR WPI: 2001-442253/47.
 DR P-PSDB: AAM41547.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX
 PS Claim 1; SEQ ID NO 4692; 10078pp; English.
 XX
 CC The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with neotropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localized neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activating/inhibiting activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemia and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 CC
 SQ Sequence 1956 BP; 474 A; 559 C; 508 G; 414 T; 1 other;
 Query Match 92.4%; Score 19.4; DB 22; Length 1956;
 Best Local Similarity 95.2%; Pred. No. 1.5;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GATACCCCCCGTCTATGAG 21
 Db 708 GATACCCCCCGTCTATGAG 728

RESULT 3
 AAI58917
 ID AAI58917 standard; cDNA; 2165 BP.
 XX
 AC AAI58917;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 1120.
 XX
 KW Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US34263.
 XX
 PR 21-JAN-2000; 2000US-0489725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.